Published in final edited form as:

Phys Med Rehabil Clin N Am. 2008 August ; 19(3): 461–viii. doi:10.1016/j.pmr.2008.04.001.

Current and Future Directions in Genomics of Amyotrophic Lateral Sclerosis

John Ravits, MD, FAAN^{a,b,c,*} and Bryan J. Traynor, MD^d

^aVirginia Mason Medical Center, 1100 Ninth Avenue, Seattle, WA 98101, USA

^bBenaroya Research Institute at Virginia Mason, 1201 Ninth Avenue, Seattle, WA 98101, USA

^cDepartment of Neurology, University of Washington School of Medicine, 1959 NE Pacific Street, Seattle, WA 98195, USA

^dNeurogenetics Branch, National Institute of Neurological Diseases and Stroke, National Institutes of Health, Building 35, Room 1A/110, 35 Convent Drive, Bethesda, MD 20892-3720, USA

Mendelian and complex genetics of familial and sporadic amyotrophic lateral sclerosis

Advances in science in the past decade have accelerated the pace of discovery in amyotrophic lateral sclerosis (ALS), especially its genomics. This article discusses the Human Genome Project, the International HapMap, and emerging microarray and micro-dissection technologies, which have set the stage for several studies seeking fundamental insights into mechanisms of diseases and targets for therapy. Directly understanding the 90% to 95% of ALS that is sporadic (SALS) as opposed to the 5% to 10% that is familial (FALS) [1,2] now seems possible.

FALS follows simple Mendelian autosomal dominant inheritance; 20% are caused by mutations in the gene encoding superoxide dismutase 1 (SOD1) at chromosome 21q22.1 [3] and the remaining 80% by unknown mutations. SALS, by contrast, is widely believed to be a complex genetic disease [2,4], which means genetic factors are important, but how much genetic factors contribute and the interplay of genetic and environmental factors are unknown. This article summarizes recent developments in genomics and efforts to apply them to understanding ALS, and highlights possible future directions of genomic research.

The Human Genome Project

The Human Genome Project was completed in 2001 and was the culmination of major advances in science, technology, and diplomacy. It was an international collaboration to sequence the 3.5 billion nucleotides in the human genome (<http://www.ncbi.nlm.nih.gov/sites/entrez?db1/4nucleotide>).

The human genome's nucleotide sequence can be divided into genic and intergenic regions. The genic regions are a minority of the code, coding for approximately 25,000 different genes. These genes are composed of promoter regions, which regulate its downstream gene; exons, which code for protein; and introns, which separate one exon from another. An average gene is 27,000 base pairs in length and consists of nine exons. In a complex

*Corresponding author: John Ravits, Benaroya Research Institute at Virginia Mason, 1201 Ninth Avenue, Seattle, WA 98101. jravits@benaroyaresearch.org.

assembly process in the nucleus, genes are transcribed into messenger RNA (mRNA), introns are removed, and exons are spliced together. The new assembly is then transported to the cytoplasm where protein is synthesized. The exact sequence in which exons are spliced together generates exon splice variants, which produce greater variations of proteins than code alone could provide, and likely plays an important role in pathogenesis of some diseases [5].

The intergenic regions, accounting for 99% of the human genome, once dismissed as “junk,” are now recognized as having critical regulatory functions [6].

The International HapMap Project

The International HapMap Project was the natural successor to the Human Genome Project [7]. It was also conducted by an international collaboration that officially began in 2002. The motivating force was to provide a practical navigation map for exploring the human genome [8]. This map is possible because interspersed throughout the genome are 11 or more million sites of predictable variation called single nucleotide polymorphisms (SNPs). Variations that appear at one SNP are often parallel or “linked” to those that appear at others, a phenomenon called linkage disequilibrium [9]. The HapMap project catalogued approximately 3.5 million SNPs across the four main ethnic groups of the human species, namely north European Caucasians, Yorubans from Nigeria, Han Chinese, and Japanese. Exact numerical measurements of linkage disequilibrium between SNPs were included. Since SNPs with high linkage disequilibrium to each other could serve as proxies for each other, a representative selection of a few hundred thousand entire genome [10]. These could then be used to search the genome.

Microarray technology

Simultaneous with the Human Genome and International HapMap Projects was the emergence of microarray technologies. These technologies simultaneously profile hundreds of thousands of DNA or RNA sequences through microchip and microbead technologies.

In this technology, a small nucleotide probe of known sequence is synthesized and attached using laser technology to a microchip at specific x- and y-coordinate locations. The resolution of the coordinates is in the range of 3 to 10 mm and millions of probes can be placed on one microchip, which can thus represent the entire genome systematically. A labeled biologic test sample, derived from DNA or RNA, is applied to the microchip and, if sequences in the biologic test sample are complementary to those on the microchip, hybridization occurs. Signal recognition software registers what sequences hybridized and the degree of the hybridization. A digital file is then generated that profiles the biologic test sample.

Many microarray platforms exist, depending on what is being sought. The most common platforms are (1) SNP arrays, which have probes designed to detect single nucleotide polymorphisms; (2) expression arrays, which have probes designed to detect genes; (3) exon arrays, which have probes designed to detect the exons that comprise genes; and (4) tiling arrays, which have probes that detect short sequences systematically throughout the genome.

Microarray technology has engendered a new investigational paradigm that is referred to as exploration or discovery because does not depend on candidate biology or prior hypotheses. Instead, a specific test condition, such as SALS, is defined and comprehensively profiled and significant patterns and hypotheses are sought [11,12]. This paradigm generates enormous data, creating major challenges to computational biology for data mining (ie, methods for searching data for meaning) and has created a new field called bioinformatics

[13]. One major challenge is the huge numbers of false leads that are generated. For example, if one microarray platform measures 550,000 items and the statistical analysis defines significance as the top 2% ($P < 0.02$), up to 11,000 findings are potentially false.

Several complex methods are emerging to better refine statistical analysis for multiple testing. One common test is the Bonferroni correction, which posits that the P value of true significance is the usual P value (ie, $P = .02$) divided by the number of hypotheses being tested (Fig. 1).

Whole genome association studies

With the International HapMap and the ability to represent the genetic variation through tagging SNPs, and the advances of the high throughput microarray technologies, one microchip became able to adequately profile genetic variation across the genome. Suddenly, studies for disease association could be genome-wide through comparing cases with a disease to controls without (case-control studies) [14].

The manner in which whole genome association (WGA) studies ascertain association is as follows. The frequencies of alleles or genotypes (combination of alleles) of each tagging SNP are ascertained in disease and non-disease populations and then statistically compared [15]. A wide variety of statistical tests to assess for association have been described. The so-called “Cochrane Armitage trend test” is a favorite because of its (1) robustness to cryptic relatedness, wherein two samples are related to each other without the knowledge of the investigator (eg, second-degree cousins); (2) deviations from Hardy-Weinberg equilibrium, which are deviations from the frequency with which an allele is in a population when steady-state is established; and (3) ability to factor risk for disease in an additive fashion, accounting for the possibility that a causative allele may have either dominant or recessive actions.

The premise of WGA studies is that disease-predisposing alleles exist with relatively high frequency. This hypothesis is often referred to as the common disease/common variant hypothesis [16]. The recent explosion in the number of published whole genome studies has also made it clear that causative genetic variants typically confer only minor to moderate risk for disease. Many associations are finding odds ratios between 1.1 and 1.4, meaning that an individual's risk for disease is increased by 10% to 40%, which seems minor for rare diseases.

Not all diseases are yielding clear results, either because of low association to disease susceptibility or because the complexity of the association, if even genetic, is created through many alleles of such low frequency that they elude detection, a hypothesis referred to as multiple rare variants hypothesis [17–19].

Whole genome association studies of sporadic amyotrophic lateral sclerosis

Four WGA studies have been published for SALS [20–23]. The first study consisted of a cohort of 276 American patients who had SALS and 275 neurologically normal American control samples [20]. Although this study identified several loci that may be important in the pathogenesis of motor neuron degeneration, none exceeded Bonferroni correction (see Fig. 1).

The second WGA study used a technique that pooled DNA to analyze a cohort of 386 patients who had SALS and 542 neurologically normal controls [21]. This study implicated

the FLJ10986 gene on chromosome 1 as being associated with increased risk for ALS (P value in replication series = 3.0×10^{-4} ; odds ratio = 1.35).

The third study, involving 461 Dutch patients who had SALS and 450 controls, identified several SNPs of interest (including one on the ITPR2 gene on chromosome 12), although none met Bonferroni threshold [22]. However, a separate but related study from this same group identified dipeptidylpeptidase 6 (DPP6) on chromosome 7 as a risk allele (P value = 3.28×10^{-6} and 5.04×10^{-8}) [24].

A fourth study involving 221 Irish patients who had ALS and 211 Irish controls identified several loci of possible importance, but none exceeded Bonferroni threshold. However, pooled analysis also identified DPP6 (combined P value = 2.53×10^{-6}), which seems to increase risk for ALS by 37% [23].

Limitations

The WGA studies have two main problems [25]. First, with the possible exception of DPP6, no loci identified in individual studies have replicated each other. Second, the observed associations seem modest. Large cohorts of several thousand cases and controls are required to have sufficient power to discriminate moderate-effect alleles, and therefore each study was underpowered in the size of the initial cohorts [26]. Efforts are underway to pool data from the various WGA studies; this combined cohort of several thousand cases and controls may have sufficient power to identify moderate effect alleles.

Whole genome expression profiling

Expression and expression arrays

In contrast to genetic association studies, which examine sites in DNA for genetic associations to disease, gene expression studies explore transcribed mRNA for gene expression in disease. These studies are therefore more rooted in biology than genetics and are sometimes referred to as functional genomics.

The basis of these studies is expression microarray technology [27]. Expression microarrays capitalize on the fact that the 3' tail of the 3' exon is polyadenylated, and therefore can serve as a tag with which to identify expressed genes. Molecular techniques are used that prime with poly-dT sequences complementary to the poly-dA tails and are thus able to separate mRNA from non-mRNA (only 1%-3% of total RNA is mRNA), amplify the mRNA, and label it [28–30].

These studies use two kinds of microarray platforms. In one, called cDNA or spotted microarrays, the probes for detecting genes are maintained from in vivo stock and spotted onto the array. In the other, called oligonucleotide microarrays, the probes are computer designed and synthesized using laser technology. This platform has the advantage of being whole genome.

Interpreting expression data is complex. Statistical methods define genes that are differentially expressed and which can then be analyzed for biologic meaning (Fig. 2) [31]. Newer methods seek enrichment of biologic processes [32–34]. Rather than differential levels of expression of any one gene, these methods seek differential expression of sets or networks of genes that work together (Fig. 3). One challenge to data interpretation is that gene function annotation is at a relatively early stage of development and is skewed in the direction of current greatest research.

Cell-specific gene expression

Microarray technologies can be used with laser micro-dissection, a computer-assisted microscopic technique that micro-dissects tissue to isolate specific cells for molecular analysis [35,36]. Through micro-dissecting and pooling a single-cell population, mRNA from the relevant cells residing in tissues with complex architecture, such as the nervous system, can be isolated and pooled [37,38].

This technique is especially useful for overcoming some main difficulties in investigating sporadic neurodegenerative diseases, such as the selectivity of the pathologic process; the variable location of pathology along the neuraxis; the reduction caused by the disease; and the low pathogenic-tonon-pathogenic signal-to-noise ratio.

ALS lends itself to study with this technique. Clinically, ALS motor neuron degeneration begins in a discrete region of the motor system and propagates outward and summates over space and time until it appears to be diffuse [39]. The outward propagation of the disease from a focus creates a gradient of neurodegeneration around the site of onset, which, because of the direct involvement of respiration, is usually still active and present at death [40]. This finding can be exploited, especially when applied to regions in early stages of degeneration with relatively early molecular events [41].

Limitations

This technique has problems and cautions. Expression profiling depends on the quality of the input, underscoring the critical importance of upstream tissue processing. Expression profiles may be incomplete and only represent 75% of a cell's expressed genes, therefore causing false-negatives (ie, mistakenly thinking that a gene that is expressed is not) [42,43]. The genes identified represent the summation of the captured neurons; even though the neuronal population is homogeneous, the cells are at various stages of development or disease and the gene expression profile is a summation of this. As with all research such as this, molecular discoveries do not differentiate between primary and secondary changes. Even if primary gene expression changes can be defined, the initial causative factors initiating this expression remain to be defined. Furthermore, the relationship between genomics and proteomics is unknown; what occurs at the gene level may not accurately reflect what is happening at the protein level [44,45].

Expression studies in amyotrophic lateral sclerosis

Several investigations have used either tissue micro-dissection [46,47] or microarray technologies [48–53] for ALS, but only one has combined them in human SALS [54], and three have combined them in transgenic models (Tables 1 and 2) [55–57].

Jiang and colleagues [54] were the first to use micro-dissection to isolate motor neurons from postmortem frozen lumbar spinal cords of patients who had SALS, and used cDNA microarrays containing 4,845 genes. They found that 3% of genes were down-regulated and were associated with cytoskeleton and axonal transport, transcription, and cell surface antigens and receptors; 1% of genes were up-regulated and included promoters for cell death pathways.

Based on these data, they further pursued expression in several genes using molecular pathology, including dynactin 1 (DCTN1), which is related to cytoskeleton/axonal transport and down-regulated; early growth response 3 (EGR3), which is related to transcription and down-regulated; death receptor 5 (DR5), which is related to cell death and up-regulated; cyclin C (CCNC), which is related to cell death and up-regulated; and acetyl-CoA

transporter (ACATN), which is related to cell death and up-regulated [58]. One key finding was that DCTN1 seems to be one of the earliest changes in motor neurons, preceding other known markers such as accumulation of phosphorylated neurofilament or ubiquitinated proteins.

Perrin and colleagues [55] were the first to report studies of transgenic mice. They studied the G93A SOD1 genotype at 60, 90, and 120 days. At the early pre-symptomatic time (60 days), 17 genes were up-regulated and 11 genes were down-regulated; approximately half the genes that were differentially expressed at the early pre-symptomatic time remained differentially expressed throughout the disease. The differentially expressed genes functioned in areas such as intermediate filaments, protein catabolism, cell communication, and regulation of cell growth, but no induction of apoptotic pathways was apparent. One gene that was particularly important because it was also identified in other models of select motor neuron degeneration was the neurofilament protein vimentin, which was up-regulated early and remained up-regulated throughout the course. When studied with immunohistochemistry, it was deposited within the cytoplasm of motor neurons in disease.

Lobsiger and colleagues [56] studied SOD1 transgenic mouse and rat models and compared a model that had active SOD1 activity (SOD1 G37R) with one that had inactive activity (SOD1 G85R) at early time points. Both models showed early induction of components of the classic complement system and the regenerative/injury response; in addition, the model that was SOD1-active also showed early dysregulation of the D/L-serine biosynthetic pathway.

Ferraiuolo and colleagues [57] also studied G93A SOD1 transgenic mouse model at 60, 90, and 120 days. At the early pre-symptomatic time (60 days), an increase was seen in transcriptional and translational functions, lipid and carbohydrate metabolism, mitochondrial pre-protein translocation, and respiratory chain function. Later in the course (90 days), up-regulation of genes involved in carbohydrate metabolism still occurred, but genes involved in transcription and mRNA processing genes begin to show down-regulation. Late in the disease course (120 days), repression of transcription and metabolic functions and up-regulation of complement system components and cyclins involved in cell-cycle regulation occurred.

Exon profiling

Variation in how exons are spliced together vastly expands the repertoire of proteins, and alterations of this phenomenon likely participate in disease pathogenesis [5]. Genome-wide profiling of exon splicing is now possible using a new microarray technology called exon arrays (Fig. 4) [59,60]. This technology capitalizes on the fact that much of the coding sequences within exons are conserved, and therefore probes to these sequences will query their individual expression, which can then be analyzed for increase or decrease.

These studies rely on a complicated scheme of random priming of mRNA to perform isolation, amplification, and labeling that differs from the expression profiles that use 30-based mRNA amplification and array platforms. Important technical, biologic, and bioinformatics issues must be addressed for this to succeed, and there is little practical experience, especially in neurologic diseases [61]. Because significant portions of SNPS identified in the WGA studies reside in introns, which if pathogenic would likely exert their effect through exon splicing, and because the nuclear factor TDP-43, a primary ubiquitinated and mislocalized protein in SALS [62,63], is a regulator of alternative splicing of exons [64], exon splice variants in SALS are important to examine.

Dark matter of the genome, noncoding RNA, and epigenetics

Vast portions of the genome code RNA transcripts that do not then code for protein (so-called “noncoding RNA”) [6]. These transcripts include microRNA, which are transcribed in nucleus and transported to cytoplasm to knock-down gene expression [65,66]; small nucleolar RNA, which reside in the nucleolus to modify genes; and small nuclear RNA, which reside in the nucleus to alter pre-mRNA splicing. Noncoding RNA are believed to be particularly important as regulators of genome function and are particularly important in the nervous system [67,68].

Various new array technologies perform high-throughput profiling of genetic codes [69-71]. These allow systematic high-density profiling of the entire genome using 25-mer probes tiled at an average of 35 base pair probe spacing, thus leaving gaps of 10 base pair. With this platform, genic (exonic and intronic) and intergenic regions can be profiled. Current limitations to this technology include amplification strategies that select sense (p) strands; the large quantities of RNA needed for the seven microarrays containing approximately 45 million oligonucleotide probes that interrogate the whole human genome; and complex bioinformatics. Soon, direct whole genome sequencing will be available. The combined efforts of computational and direct biology is critical to exploring this area [72-75].

Epigenetics is the study of gene expression and regulation that occur separate from change in DNA sequence [76,77]. One important area is DNA methylation, which mutes gene expression through adding a methyl group to the 5-position of the cytosine ring. New technologies are emerging for high-throughput profiling of methylation changes, thus profiling the so called “methylome,” and this could reveal fundamental aspects of ALS pathogenesis [69,70].

Summary

High-resolution exploration of the genome is now possible and can be applied to understanding SALS biology, including exploration of the DNA, which might provide clues to SALS susceptibility, and of the RNA, which might provide clues of the SALS gene expression and molecular pathogenesis. Both genic regions, including introns and exons and their splicing, and non-genic regions, including structural, regulatory, and other elements, can be explored. Micro and nanotechnologies, informatics, and computational biology could potentially lead to a deep understanding of molecular pathogenesis and identify targets for therapy.

Acknowledgments

Dr. Traynor's work was supported by the intramural programmes of the National Institute on Aging (NIA) and the National Institute on Neurological Disorders and Stroke (NINDS). Dr. Ravits' work was supported by the National Institute of Health (R21 NS05173801A1), Department of Defense (USAMRMC Proposal #06054001), the Moyer Foundation, the Juniper Foundation, the Benaroya Foundation, and private philanthropists.

References

- [1]. Figlewicz DA, Orrell RW. The genetics of motor neuron disease. Amyotroph Lateral Scler Other Motor Neuron Disord. 2003; 4:225–31. [PubMed: 14753656]
- [2]. Valdmantis PN, Rouleau GA. Genetics of familial amyotrophic lateral sclerosis. Neurology. 2008; 70:144–52. [PubMed: 18180444]
- [3]. Andersen PM. Amyotrophic lateral sclerosis associated with mutations in the Cu Zn superoxide dismutase gene. Curr Neurol Neurosci Rep. 2006; 6:37–46.
- [4]. Simpson CL, Al-Chalabi A. Amyotrophic lateral sclerosis as a complex genetic disease. Biochim Biophys Acta. 2006; 1762:973–85. [PubMed: 16973338]

- [5]. Novoyatleva T, Tang Y, Rafalska I, et al. Pre-mRNA missplicing as a cause of human disease. *Prog Mol Subcell Biol.* 2006; 44:27–46. [PubMed: 17076263]
- [6]. Johnson JM, Edwards S, Shoemaker D, et al. Dark matter in the genome: evidence of widespread transcription detected by microarray tiling experiments. *Trends Genet.* 2005; 21(2):93–102. [PubMed: 15661355]
- [7]. The International HapMap Consortium. A haplotype of the human genome. *Nature.* 2005; 437:1299–320. [PubMed: 16255080]
- [8]. Barnes MR. Navigating the HapMap. *Brief Bioinform.* 2006; 7:211–24. [PubMed: 16877472]
- [9]. Hartyl, DL.; Clark, AG. Principles of population genetics. 4th edition. Sinauer Associates, Inc.; Sunderland (MA): 2007. p. 652
- [10]. Abecasis GR, Ghosh D, Nichols TE. Linkagedisequilibrium:ancienthistorydrivesthenew genetics. *Hum Hered.* 2005; 59:118–24. [PubMed: 15838181]
- [11]. Brown PO, Botstein D. Exploring then ewworldof the genome with DNA micro arrays. *Nat Genet.* 1999; 21(1 Suppl):33–7. [PubMed: 9915498]
- [12]. Goodman L. Hypothesis-limited research. *Genome Res.* 1999; 9:673–4. [PubMed: 10447502]
- [13]. Barnes, MR., editor. Bioinformatics for geneticists. A bioinformatics primer for the analysis of genetic data. 2nd edition. Wiley and Sons; West Sussex (England): 2007. p. 554
- [14]. Hirschhorn JM, Daly MJ. Genome-wide association studies for complex diseases and complex traits. *Nat Rev Genet.* 2005; 6:95–108. [PubMed: 15716906]
- [15]. Balding DJ. A tutorial on statistical methods for population association studies. *Nat Rev Genet.* 2006; 7:781–91. [PubMed: 16983374]
- [16]. Reich DE, Lander ES. Ontheallelicspectrumofhumandisease. *Trends Genet.* 2001; 17:502–10. [PubMed: 11525833]
- [17]. Pritchard JK. Arerarevariantsresponsibleforsusceptibilityto complexdiseases? *AmJHum Genet.* 2001; 69:124–37.
- [18]. Terwilliger JD, Hiekkalinna T. An utter refutation of the fundamental theorem of the HapMap. *Eur J Hum Genet.* 2006; 14:426–37. [PubMed: 16479260]
- [19]. Wright AF, Hastie ND. Complex genetic diseases: controversy over the Croesus code. *Genome Biol.* 2001; 2(8)
- [20]. Schymick J, Scholz SW, Fung H-C, et al. Genome-wide genotyping in amyotrophic lateral sclerosis and neurologically normal controls. *Lancet Neurol.* 2007; 6(4):322–8. [PubMed: 17362836]
- [21]. Dunckley T, Huentelman MJ, Craig DW, et al. Whole-genome analysis of sporadic amyotrophic lateral sclerosis. *NEJM.* 2007; 357:1–14. [PubMed: 17611201]
- [22]. van Es MA, van Vught PW, Blauw HW, et al. ITPR2 as a susceptibility gene in sporadic amyotrophic lateral sclerosis: a genome-wide association study. *Lancet Neurol.* 2007; 6(10):869–77. [PubMed: 17827064]
- [23]. Cronin S, Berger S, Ding J, et al. A genome-wide association study of sporadic ALS in a homogenous Irish population. *Hum Mol Genet.* 2008; 17:768–74. [PubMed: 18057069]
- [24]. van Es MA, van Vught PW, Blauw HM, et al. Genetic variation in DPP6 is associated with susceptibility to amyotrophic lateral sclerosis. *Nat Genet.* 2008; 40(1):29–31. [PubMed: 18084291]
- [25]. Garber K. The elusive ALS genes. *Science.* 2008; 319(5859):20. [PubMed: 18174402]
- [26]. Skol AD, Scott LJ, Abecasis GR, et al. Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. *Nat Genet.* 2006; 38:209–13. [PubMed: 16415888]
- [27]. Kirby J, Heath PR, Shaw PJ, et al. Geneexpressionassays. *AdvClinChem.* 2007; 44:247–92.
- [28]. Li J, Schwartz SM, Bumgarner RE. RNA amplification, fidelity and reproducibility of expression profiling. *Comptes Rendus Biol.* 2003; 326:1021–30.
- [29]. Ginsberg SD, Che S. RNA amplification in brain tissues. *Neurochem Res.* 2002; 27:981–92. [PubMed: 12462399]

- [30]. Kelz MB, Dent GW, Therianos S, et al. Single-cell antisense RNA amplification and micro-array analysis as a tool for studying neurological degeneration and restoration. *Sci Aging Knowledge Environ.* 2002; 1:1–10.
- [31]. Speed, T., editor. *Statistical analysis of gene expression microarray data*. Boca Raton (FL):
- [32]. Subramanian A, Tamayo P, Mootha V, et al. Gene set enrichment analysis: a knowledgebased approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A.* 2005; 102:15545–50. [PubMed: 16199517]
- [33]. Bild A, Febbo PG. Application of a priori established gene sets to discover biologically important differential expression in microarray data. *Proc Natl Acad Sci U S A.* 2005; 107:15278–9. [PubMed: 16230612]
- [34]. Efron, B.; Tibshirani, R. On testing the significance of sets of genes. Available at: <http://www-stat.stanford.edu/whitbs/ftp/GSA.pdf>
- [35]. Emmert-Buck MR, Bonner RF, Smith PD, et al. Laser capture microdissection. *Science.* 1996; 274:998–1001. [PubMed: 8875945]
- [36]. Bonner RF, Emmert-Buck KM, Cole K, et al. Laser capture microdissection: molecular analysis of tissue. *Science.* 1997; 278:1481–3. [PubMed: 9411767]
- [37]. Nisenbaum LK. The ultimate chipshot: can microarray technology deliver for neuroscience? *Genes Brain Behav.* 2002; 1:27–34. [PubMed: 12886947]
- [38]. Mills JC, Roth KA, Cagan RL, et al. DNA microarrays and beyond: completing the journey from tissue to cell. *Nat Cell Biol.* 2001; 3:E175–8. [PubMed: 11483971]
- [39]. Ravits J, Paul P, Jorg C. Focality of upper and lower motor neuron degeneration at the clinical onset of ALS. *Neurology.* 2007; 68:1571–5. [PubMed: 17485643]
- [40]. Ravits J, Laurie P, Fan Y, et al. Implications of ALS focality: rostral-caudal distribution of lower motor neuron loss postmortem. *Neurology.* 2007; 68:1576–82. [PubMed: 17485644]
- [41]. Ravits J, Laurie P, Stone B. Amyotrophic lateral sclerosis microgenomics. *Phys Med Rehabil Clin N Am.* 2005; 16(4):909–24. [PubMed: 16214051]
- [42]. Kacharmina JE, Crino PB, Eberwine J. Preparation of cDNA from single cells and sub cellular regions. *Methods Enzymol.* 1999; 303:3–18. [PubMed: 10349635]
- [43]. Crino PB, Trojanowski JQ, Dichter MA, et al. Embryonic neuronal markers in tuberous sclerosis: single-cell molecular pathology. *Proc Natl Acad Sci USA.* 1996; 93:14152–7. [PubMed: 8943076]
- [44]. Humphery-Smith I, Cordwell SJ, Blackstock WP. Proteome research: complementarity and limitations with respect to the RNA and DNA worlds. *Electrophoresis.* 1997; 18:1217–42. [PubMed: 9298643]
- [45]. Greenbaum D, Colangelo C, Williams K, et al. Comparing protein abundance and mRNA expression levels on a genomic scale. *Genome Biol.* 2003; 4(9):117. [PubMed: 12952525]
- [46]. Heath PR, Tomkins J, Ince PG, et al. Quantitative assessment of AMPA receptor mRNA in human spinal motor neurons isolated by laser capture microdissection. *Neuroreport.* 2002; 13(14):1753–7. [PubMed: 12395117]
- [47]. Mawrin C, Kirches E, Dietzmann K. Single-cell analysis of mtDNA in amyotrophic lateral sclerosis: towards the characterization of individual neurons in neurodegenerative disorders. *Pathol Res Pract.* 2003; 199:415–8. [PubMed: 12924443]
- [48]. Malaspina A, Kaushik N, deBelleroche J. Differential expression of 14 genes in amyotrophic lateral sclerosis spinal cord detected using gridded cDNA arrays. *J Neurochem.* 2001; 77:132–45. [PubMed: 11279269]
- [49]. Ishigaki S, Niwa J, Ando Y, et al. Differentially expressed genes in sporadic amyotrophic lateral sclerosis spinal cords—screening by molecular indexing and subsequent cDNA micro-array analysis. *FEBS Lett.* 2002; 531:354–8. [PubMed: 12417341]
- [50]. Dangond F, Hwang D, Camelo S, et al. Molecular signature of late-stage human ALS revealed by expression profiling of postmortem spinal cord gray matter. *Physiol Genomics.* 2004; 16:229–39. [PubMed: 14645737]
- [51]. Olsen MK, Roberds SL, Ellerbrock BR, et al. Disease mechanisms revealed by transcription profiling in SOD1-G93A transgenic mouse spinal cord. *Ann Neurol.* 2001; 50:730–40. [PubMed: 11761470]

- [52]. Yoshihara T, Ishigaki S, Yamamoto M, et al. Differential expression of inflammation and apoptosis-related genes in spinal cords of a mutant SOD1 transgenic mouse model of familial amyotrophic lateral sclerosis. *J Neurochem.* 2002; 80:158–67. [PubMed: 11796754]
- [53]. Hensley K, Floyd RA, Gordon B, et al. Temporal patterns of cytokine and apoptosis-related gene expression in spinal cords of the G93A-SOD1 mouse model of amyotrophic lateral sclerosis. *J Neurochem.* 2002; 82:365–74. [PubMed: 12124437]
- [54]. Jiang Y-M, Yamamoto M, Koyayashi Y, et al. Gene expression profile of spinal motor neurons in sporadic amyotrophic lateral sclerosis. *Ann Neurol.* 2005; 57:236–51. [PubMed: 15668976]
- [55]. Perrin F, Boisset G, Docquier M, et al. No widespread induction of cell death genes occurs in pure motoneurons in an amyotrophic lateral sclerosis mouse model. *Hum Mol Genet.* 2005; 14:3309–20. [PubMed: 16192287]
- [56]. Lobsiger CS, Boillee S, Cleveland DW. Toxicity from different SOD1 mutants dysregulates the complement system and the neuronal regenerative response in ALS motor neurons. *Proc Natl Acad Sci U S A.* 2007; 104:7319–26. [PubMed: 17463094]
- [57]. Ferraiuolo L, Heath PR, Holden H, et al. Microarray analysis of the cellular pathways involved in the adaptation to and progression of motor neuron injury in the SOD1 G93A mouse model of familial ALS. *J Neurosci.* 2007; 27:9201–19. [PubMed: 17715356]
- [58]. Jiang YM, Yamamoto M, Tanaka F, et al. Gene expression specifically detected in motor neurons (dynactin 1, early growth response 3, acetyl-CoA transporter, death receptor 5, and cyclin C) differentially correlate to pathologic markers in sporadic amyotrophic lateral sclerosis. *J Neurophol Exp Neurol.* 2007; 66:617–27.
- [59]. Cuperlovic-Culf M, Belacel N, Culf AS, et al. Microarray analysis of alternative splicing. *OMICS.* 2006; 10(3):344–57. [PubMed: 17069512]
- [60]. Clark TA, Schweitzer AC, Chen TX, et al. Discovery of tissue-specific exons using comprehensive human exon microarrays. *Genome Biol.* 2007; 8(4):R64, 1–16. [PubMed: 17456239]
- [61]. Heinzen EL, Yoon W, Weale ME, et al. Alternative ion channel splicing in mesial temporal lobe epilepsy and Alzheimer's disease. *Genome Biol.* 2007; 8(3):R32, 1–18. [PubMed: 17343748]
- [62]. Newmann M, Sampathu DM, Kwong LK, et al. Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science.* 2006; 314(5796):130–3. [PubMed: 17023659]
- [63]. Arai T, Hasegawa LM, Akiyama H, et al. TDP-43 is a component of ubiquitin-positive tau-negative inclusions in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Biochem Biophys Res Commun.* 2006; 351:602–11. [PubMed: 17084815]
- [64]. Buratti E, Dork T, Zuccato E, et al. Nuclear factor TDP-43 and SR proteins promote *in vitro* and *in vivo* CFTR exon 9 skipping. *EMBO J.* 2001; 20(7):1774–84. [PubMed: 11285240]
- [65]. Krichevsky AM. MicroRNA profiling: from dark matter to white matter, or identifying new players in neurobiology. *Scientific World Journal.* 2007; 7(S2):155–66. [PubMed: 17982589]
- [66]. Kim VN, Nam J-W. Genomics of microRNA. *Trends Genet.* 2005; 22:165–73. [PubMed: 16446010]
- [67]. Mehler MF, Mattick JS. Non-coding RNAs in the nervous system. *J Physiol.* 2006; 575:333–41. [PubMed: 16809366]
- [68]. Nelson PT, Keller JN. RNA in brain disease: no longer just “the messenger in the middle”. *J Neuropathol Exp Neurol.* 2007; 66:461–8. [PubMed: 17549006]
- [69]. Mockler TC, Ecker JR. Applications of DNA tiling arrays for whole-genome analysis. *Genomics.* 2005; 85:1–15. [PubMed: 15607417]
- [70]. Hoheisel J. Microarray technology: beyond transcript profiling and genotype analysis. *Nat Rev Genet.* 2006; 7:200–10. [PubMed: 16485019]
- [71]. Cowell JK, Hawthorn L. The application of microarray technology to the analysis of the cancer genome. *Curr Mol Med.* 2007; 7:103–20. [PubMed: 17311536]
- [72]. Chaudhuri K, Chatterjee R. MicroRNA detection and target prediction: integration of computational and experimental approaches. *DNA Cell Biol.* 2007; 26:321–37. [PubMed: 17504028]

- [73]. Berezikow E, Cuppen E, Plasterek RH. Approaches to microRNA discovery. *Nat Genet.* 2006; 38:52–7.
- [74]. Zhang B, Pan X, Wang Q, et al. Computational identification of microRNAs and their targets. *Comput Biol Chem.* 2006; 30:395–407. [PubMed: 17123865]
- [75]. Huang JC, Morris QD, Frey BJ. Bayesian inference of microRNA targets from sequence and expression data. *J Comput Biol.* 2007; 14:550–63. [PubMed: 17683260]
- [76]. Saetrom P, ve O Jr, Rossi JJ. Epigenetics and microRNA. *PediatrRes.* 2007; 61:17R–23R.
- [77]. Chuang JC, Jones PA. Epigenetics and microRNA. *Pediatr Res.* 2007; 61:24R–9R.

Fig. 1.

Whole genome SNP association study. The plot displays the degree of association of particular SNPs with SALS: the x-axis displays SNPs chromosomal position and the y-axis shows $-\log_{10}(P)$ value obtained by allelic association test). The red line represents threshold for significance after Bonferroni correction for multiple testing. (Data from Schymick J, Scholz SW, Fung H-C et al. Genome-wide genotyping in amyotrophic lateral sclerosis and neurologically normal controls. *Lancet Neurol* 2007;6(4):322–8.)

Fig. 2.

Differential gene expression. The plots show gene expression in motor neurons and anterior horns in lumbar spinal cords of SALS. The motor neurons were isolated from the surrounding anterior horn using a technique called laser capture microdissection, which allows histologic sections to be dissected with a specialized microscope. Gene expressions were performed using a technology called whole genome oligonucleotide microarray, which measure expression of each gene. The top plots compare anterior horns to motor neurons in control (left) and SALS (right); the bottom plots compare SALS to control in the neurons (left) and anterior horn (right.) (Data from John Ravits, MD, unpublished.)

Fig. 3.

Differential gene set expression. The plots show the neuron migration pathway of 54 genes that seems to be up-regulated in SALS motor neurons. The left plot shows what is called a heat map, which is a way of visualizing patterns. In the example, the groups of interest are in columns (control on the left and SALS on the right) and the genes in the pathway of interest are in rows (the top rows are the most differentiated between the groups). The right had plot shows the overall up-regulation (shift to the right) in the expression levels of these genes. (Data from John Ravits, MD, unpublished.)

Fig.4.

Exon profiling and splicing. The plots show the expression levels of the exons comprising the TARDBP gene, here chosen to illustrate what is possible. They are taken from a technology called whole genome exon microarrays, which are able to measure each exon of each gene throughout the genome. The data is then processed by specialized computer software. The samples are obtained from motor neurons that were isolated by laser capture microdissection. The red graph is SALS and the blue graph is control. TARDBP is the gene coding for a protein called TDP-43, which is mislocalized in SALS. One of its functions is exon splicing. (Data from Stuart Rabin, PhD and John Ravits, MD, unpublished.)

\$watermark-text

\$watermark-text

\$watermark-text

Table 1

Published genome wide association studies in amyotrophic lateral sclerosis

Population	Cases	Controls	Method	Gene described (P value)	Reference
US	276	271	HumanHap550K (Illumina)	No clearly associated gene	Schymick et al [20]
US	386	542	Human Mapping 500K (Affymetrix) p HumanHap300 (Illumina)	FLJ10986 (pooled P = 3.0×10^{-4})	Dunkley et al [21]
Dutch	461	450	HumanHap300K (Illumina)	ITPR2 (3.5×10^{-6}); DPP6 (pooled P = 5.0×10^{-8})	van Es et al [22,24]
Irish	222	217	HumanHap550k (Illumina)	No clearly associated gene; DPP6 (pooled P = 2.53×10^{-6})	Cronin et al [23]

DNA samples were pooled and then results between SNP typing platforms were pooled.

Table 2

Microarray studies of microdissected motor neurons

Tissue	Microarray	Time points	Controls	Method of analysis	Major findings	Reference
SALS	cDNA microarray of 4845 selected genes	NA	Non-SALS	Differential gene expression (O3-fold change)	Up-regulated promoters for cell death pathways; downregulated cytoskeleton and axonal transport, transcription, and cell surface antigens and receptors	Jiang et al [54]
SOD1 G93A Tg mice	Whole genome oligonucleotide microarray	60, 90, and 120 days	G93A non-Tg littermates	Differential gene expression (O1.5-fold change)	Early and persistent dysregulation of intermediate filaments, protein catabolism, cell communication, and regulation of cell growth	Perrin et al [55]
SOD1 G37R and G85R Tg mice; G93A Tg rats	Whole genome oligonucleotide microarray	8 and 15 weeks	SOD1 WT Tg mice+ rats	Differential gene expression (R1.5-fold change)	Early dysregulation of the D/L-serine biosynthetic pathway, induction of components of the classic complement system and of the regenerative/injury response	Lobsiger et al [56]
SOD1 G93A Tg mice	Whole genome oligonucleotide microarray	60, 90, and 120 days	G93A non-Tg littermates; SOD1 WT Tg mice; wild-type SOD1 transgenic mice littermates	Differential gene expression (O2.0-fold change)	Early dysregulation in transcriptional and translational functions, lipid, and carbohydrate metabolism, mitochondrial preprotein translocation, and respiratory chain function	Ferraiuolo et al [57]